

Further Corroboration of Equivalence of Two χ^2 Forms*

X.H.Mo¹; [†]

1 (Institute of High Energy Physics, CAS, Beijing 100039, China)

July 26, 2005

Abstract The equivalence of two χ^2 forms for linear function fit is proved by mathematical calculation. The simplified R-value measurements are quoted to test the conclusion quantitatively.

Key words equivalence, χ^2 form, linear function, R-value measurement

1 Introduction

The covariance matrix is usually used to construct the χ^2 estimator for correlative data which is to be minimized to acquire best estimates for parameters interested ^[1]. It is frequently the case that experimental data are affected by overall systematic error, such as the error of luminosity or efficiency in scan experiment. As a typical example without lossing generality, let n_i be the selected number of event for certain final state at the i -th energy point, and ϵ the corresponding efficiency for all measured points¹. Then the number of measured event is calculated as

$$y_i = \frac{n_i}{\epsilon} .$$

Since all y_i contain the same ϵ , they are correlated and ϵ here could be treated as a normalization factor. Under such case, the $n \times n$ covariance matrix \mathbf{V} for n measurements could be constructed as follows: the diagonal elements are given by

$$v_{ii} = \sigma_i^2 + y_i^2 \sigma_f^2 ,$$

*supported by National Natural Science Foundation of China (10491303) and 100 Talents Programme of CAS (U-25).

[†]E-mail: moxh@mail.ihep.ac.cn

¹This is a fairly good approximation for the efficiency at continuum region, where no resonance exists.

where σ_i is the statistic uncertainty of n_i and σ_f is the relative uncertainty of ϵ , the quantity $y_i\sigma_f$ is the normalization uncertainty due to factor ϵ for variable y_i . The correlation between data points i and j contributes to the off-diagonal matrix element v_{ij} , which is depicted by the product of two normalization uncertainties, that is

$$v_{ij} = y_i y_j \sigma_f^2 .$$

Explicitly, the covariance matrix has the form²

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 + y_1^2 \sigma_f^2 & y_1 y_2 \sigma_f^2 & \cdots & y_1 y_n \sigma_f^2 \\ y_2 y_1 \sigma_f^2 & \sigma_2^2 + y_2^2 \sigma_f^2 & \cdots & y_2 y_n \sigma_f^2 \\ \vdots & \vdots & \ddots & \vdots \\ y_n y_1 \sigma_f^2 & y_n y_2 \sigma_f^2 & \cdots & \sigma_n^2 + y_n^2 \sigma_f^2 \end{pmatrix} . \quad (1)$$

The expression to be minimized is then³

$$\chi_M^2 = \eta^\tau \mathbf{V}^{-1} \eta, \quad (2)$$

where

$$\eta = \begin{pmatrix} y_1 - k_1 \\ y_2 - k_2 \\ \vdots \\ y_n - k_n \end{pmatrix}$$

is the vector of the residuals between experimental observations y_i and theoretical expectation k_i .

Apart from **matrix method**, an alternative way to handle correlation problem is called **factor method** [3, 4]. A normalization factor f is introduced and is to be fitted as a free parameters to take the correlations into account

$$\chi_f^2 = \sum_i \frac{(f y_i - k_i)^2}{\sigma_i^2} + \frac{(f - 1)^2}{\sigma_f^2} . \quad (3)$$

where σ_i and σ_f have the same meaning as before.

Actually, we can comprehend the Eq.(3) from an experimental point of view. We treat the efficiency ϵ also as a measured variable and introduce a fitting parameter ϵ_0 as its expected value, then we construct a chisquare form for $n+1$ uncorrelated measurements as follows:

$$\chi^2 = \sum_i \frac{(y'_i - k_i)^2}{\sigma_i^2} + \frac{(\epsilon - \epsilon_0)^2}{(\delta\epsilon)^2} , \quad (4)$$

²According to the error propagation formula, we can derive the covariance matrix fair formally, more details can be found in Ref. [5].

³In this paper, chi-square minimization is adopted to obtain the best estimated value. For experimental data minimization, the MINUIT package is used. The detail explanation about chi-square minimization technique and MINUIT package can be found in Ref. [2].

where $\delta\epsilon$ is the efficiency uncertainty, and $y'_i \equiv n_i/\epsilon_0$, which can be rewritten as

$$y'_i = \frac{n_i}{\epsilon_0} = \frac{n_i}{\epsilon_i} \frac{\epsilon_i}{\epsilon_0} = f y_i .$$

Here we define the $f = \epsilon/\epsilon_0$. At the same time, we change the last term form of Eq. (4), viz.

$$\frac{(\epsilon/\epsilon_0 - \epsilon_0/\epsilon_0)^2}{(\delta\epsilon/\epsilon_0)^2} = \frac{(f - 1)^2}{(\delta\epsilon/\epsilon_0)^2} .$$

Experimentally, ϵ_0 could be replaced by the measurable quantity ϵ , so $\delta\epsilon/\epsilon_0 \simeq \delta\epsilon/\epsilon = \sigma_f$. Under such case, we immediately recover Eq.(3).

Now there are two χ^2 forms, as expressed in Eqs.(2) and (3), which have been used to deal with the correlated data. As a matter of fact, the equivalence between these two χ^2 forms was first discussed in detail by D'Agostini in Ref. [5]. At the same time, D'Agostini pointed out the biasness of these two estimators. Nevertheless, the D'Agostini's study was restricted within the two-variate and constant fitting case. It is natural to ask

1. is the equivalence of two χ^2 forms a general conclusion?
2. does the biasness of two χ^2 forms exist always?

As to the first question, the equivalent conclusion has been expanded to the multi-variate and constant fitting case [6, 7]. Recently, effects have been made in order to extend equivalent conclusion to the non-constant fitting case. As the first step, the linear dependent case is considered in this paper, and the equivalence is proved strictly based on mathematical calculation. So far as the second question is concerned, the following study indicates for linear function fit, the biasness still exist. Fortunately, by virtue of factor method, the biasness due to fit can be corrected through the fitted factor.

Besides the experimental requirement, statistically, the analytical results are appreciated which could provide some qualitatively understandings towards more complicated problems and some new clues for further exploration. In the following sections, we first work out the minimization values of two χ^2 forms, exactly the same forms of which lead to equivalent conclusion directly. Then some simplified experiment results are quoted to test the equivalence quantitatively.

2 Equivalence Proof

Eqs. (2) and (3) are the two χ^2 forms we will study. Where k_i linearly depends on i , or

$$k_i = \alpha \cdot i + \beta .$$

However in actual experiment, the physical attribute is often indicated by a physical variable instead of the sequence of experiment. For the i -th experiment at energy $E_{cm} =$

x_i , we usually use x_i instead of i to denote such experiment. So more practically, we write the linear form of k_i as

$$k_i = \alpha \cdot x_i + \beta \quad . \quad (5)$$

With above expression, Eq. (2) reads explicitly

$$\chi_M^2 = \sum_{i=1}^n \sum_{j=1}^n [y_i - (\alpha x_i + \beta)] \cdot \lambda_{ij} \cdot [y_j - (\alpha x_j + \beta)] \quad . \quad (6)$$

Here, the symbol $(\mathbf{V}^{-1})_{ij}$ has been changed into λ_{ij} for the convenience of notation. For the factor method, chi-square has the comparatively simplified form:

$$\chi_f^2 = \sum_{i=1}^n \frac{[f y_i - (\alpha x_i + \beta)]^2}{\sigma_i^2} + \frac{(f-1)^2}{\sigma_f^2} \quad . \quad (7)$$

In above equations, subscripts M and f indicate the matrix and the factor methods, respectively.

2.1 Expectation and variance from covariance matrix method

For briefness, we introduce matrix notation

$$\mathbf{T} = \frac{1}{2} \begin{pmatrix} \frac{\partial^2 \chi_M^2}{\partial \alpha \partial \alpha} & \frac{\partial^2 \chi_M^2}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \chi_M^2}{\partial \beta \partial \alpha} & \frac{\partial^2 \chi_M^2}{\partial \beta \partial \beta} \end{pmatrix} \equiv \begin{pmatrix} T_{xx} & T_x \\ T_x & T_0 \end{pmatrix} \quad .$$

According to the minimization condition

$$\begin{cases} \frac{\partial \chi_M^2}{\partial \alpha} = 0, \\ \frac{\partial \chi_M^2}{\partial \beta} = 0, \end{cases}$$

we have

$$\mathbf{T} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} T_{xy} \\ T_y \end{pmatrix} \quad . \quad (8)$$

In above equations, we have defined

$$\begin{aligned} T_{xx} &= \sum_{ij} x_i \lambda_{ij} x_j \quad , \quad T_x = \sum_{ij} x_i \lambda_{ij} = \sum_{ij} \lambda_{ij} x_j \quad , \\ T_{xy} &= \sum_{ij} x_i \lambda_{ij} y_j \quad T_y = \sum_{ij} \lambda_{ij} y_j \quad , \quad \text{and} \quad T_0 = \sum_{ij} \lambda_{ij} \quad . \end{aligned} \quad (9)$$

Solving Eq. (8), we obtain

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \mathbf{T}^{-1} \begin{pmatrix} T_{xy} \\ T_y \end{pmatrix} = \frac{1}{D_{\mathbf{T}}} \cdot \begin{pmatrix} T_0 T_{xy} - T_x T_y \\ T_{xx} T_y - T_x T_{xy} \end{pmatrix} \quad ,$$

and covariance of α and β

$$\begin{pmatrix} \sigma_{\hat{\alpha}}^2 \\ \sigma_{\hat{\beta}}^2 \end{pmatrix} = \frac{1}{D_{\mathbf{T}}} \cdot \begin{pmatrix} T_0 \\ T_{xx} \end{pmatrix} ,$$

with

$$D_{\mathbf{T}} \equiv |\mathbf{T}| = T_{xx}T_0 - T_xT_x . \quad (10)$$

By virtue of definitions in Eqs. (9) and (10), and formulas (A. 2) and (A. 3) in appendix, we can work out the following results

$$\begin{cases} \hat{\alpha} = \sum_{ij} \frac{x_i y_i - x_i y_j}{\sigma_i^2 \sigma_j^2} \Big/ (S \cdot D_{\mathbf{T}}) , \\ \hat{\beta} = \sum_{ij} \frac{x_i^2 y_j - x_i x_j y_j}{\sigma_i^2 \sigma_j^2} \Big/ (S \cdot D_{\mathbf{T}}) ; \end{cases} \quad (11)$$

and

$$\begin{cases} \sigma_{\hat{\alpha}}^2 = \left(\sum_i \frac{1}{\sigma_i^2} + \sigma_f^2 \cdot \sum_{ij} \frac{y_i^2 - y_i y_j}{\sigma_i^2 \sigma_j^2} \right) \Big/ (S \cdot D_{\mathbf{T}}) , \\ \sigma_{\hat{\beta}}^2 = \left(\sum_i \frac{x_i^2}{\sigma_i^2} + \sigma_f^2 \cdot \sum_{ij} \frac{x_i^2 y_j^2 - x_i y_i x_j y_j}{\sigma_i^2 \sigma_j^2} \right) \Big/ (S \cdot D_{\mathbf{T}}) , \end{cases} \quad (12)$$

where

$$D_{\mathbf{T}} = \frac{1}{S} \cdot \left[\sum_{ij} \frac{x_i^2 - x_i x_j}{\sigma_i^2 \sigma_j^2} + \sigma_f^2 \cdot \sum_{ijk} \frac{(x_i^2 - x_i x_j) y_k^2 - x_i^2 y_j y_k - x_i y_i x_j y_j + 2 x_i y_i x_j y_k}{\sigma_i^2 \sigma_j^2 \sigma_k^2} \right] . \quad (13)$$

2.2 Expectation and variance from factor method

For briefness, we introduce matrix notation

$$\mathbf{\Lambda} = \frac{1}{2} \begin{pmatrix} \frac{\partial^2 \chi_f^2}{\partial \alpha \partial \alpha} & \frac{\partial^2 \chi_f^2}{\partial \alpha \partial \beta} & \frac{\partial^2 \chi_f^2}{\partial \alpha \partial f} \\ \frac{\partial^2 \chi_f^2}{\partial \beta \partial \alpha} & \frac{\partial^2 \chi_f^2}{\partial \beta \partial \beta} & \frac{\partial^2 \chi_f^2}{\partial \beta \partial f} \\ \frac{\partial^2 \chi_f^2}{\partial f \partial \alpha} & \frac{\partial^2 \chi_f^2}{\partial f \partial \beta} & \frac{\partial^2 \chi_f^2}{\partial f \partial f} \end{pmatrix} \equiv \begin{pmatrix} \mathcal{A} & \mathcal{D} & -\mathcal{E} \\ \mathcal{D} & \mathcal{B} & -\mathcal{F} \\ -\mathcal{E} & -\mathcal{F} & \mathcal{C} \end{pmatrix} .$$

According to the minimization condition

$$\begin{cases} \frac{\partial \chi_f^2}{\partial \alpha} = 0, \\ \frac{\partial \chi_f^2}{\partial \beta} = 0, \\ \frac{\partial \chi_f^2}{\partial f} = 0, \end{cases}$$

we have

$$\mathbf{\Lambda} \begin{pmatrix} \alpha \\ \beta \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \Delta \end{pmatrix} . \quad (14)$$

In above equations, we have defined

$$\begin{aligned} \mathcal{A} &= \sum_i \frac{x_i^2}{\sigma_i^2} , \quad \mathcal{B} = \sum_i \frac{1}{\sigma_i^2} , \quad \mathcal{C} = \frac{1}{\sigma_f^2} + \sum_i \frac{y_i^2}{\sigma_i^2} , \\ \mathcal{D} &= \sum_i \frac{x_i}{\sigma_i^2} , \quad \mathcal{E} = \sum_i \frac{x_i y_i}{\sigma_i^2} , \quad \mathcal{F} = \sum_i \frac{y_i}{\sigma_i^2} , \quad \text{and} \quad \Delta = \frac{1}{\sigma_f^2} . \end{aligned} \quad (15)$$

Solving Eq. (14), and utilizing the definitions in Eq. (15), we obtain

$$\begin{cases} \hat{\alpha} = \sum_{ij} \frac{x_i y_i - x_i y_j}{\sigma_i^2 \sigma_j^2} \Big/ (\sigma_f^2 \cdot D_{\mathbf{\Lambda}}) , \\ \hat{\beta} = \sum_{ij} \frac{x_i^2 y_j - x_i x_j y_j}{\sigma_i^2 \sigma_j^2} \Big/ (\sigma_f^2 \cdot D_{\mathbf{\Lambda}}) , \\ \hat{f} = \sum_{ij} \frac{x_i^2 - x_i x_j}{\sigma_i^2 \sigma_j^2} \Big/ (\sigma_f^2 \cdot D_{\mathbf{\Lambda}}) ; \end{cases} \quad (16)$$

and the corresponding covariance of above quantities

$$\begin{cases} \sigma_{\hat{\alpha}}^2 = \left(\sum_i \frac{1}{\sigma_i^2} + \sigma_f^2 \cdot \sum_{ij} \frac{y_i^2 - y_i y_j}{\sigma_i^2 \sigma_j^2} \right) \Big/ (\sigma_f^2 \cdot D_{\mathbf{\Lambda}}) , \\ \sigma_{\hat{\beta}}^2 = \left(\sum_i \frac{x_i^2}{\sigma_i^2} + \sigma_f^2 \cdot \sum_{ij} \frac{x_i^2 y_j^2 - x_i y_i x_j y_j}{\sigma_i^2 \sigma_j^2} \right) \Big/ (\sigma_f^2 \cdot D_{\mathbf{\Lambda}}) , \\ \sigma_{\hat{f}}^2 = \left(\sum_{ij} \frac{x_i^2 - x_i x_j}{\sigma_i^2 \sigma_j^2} \right) \Big/ D_{\mathbf{\Lambda}} ; \end{cases} \quad (17)$$

where

$$\begin{aligned} D_{\mathbf{\Lambda}} &\equiv |\mathbf{\Lambda}| = \mathcal{A}\mathcal{B}\mathcal{C} - \mathcal{C}\mathcal{D}^2 - \mathcal{B}\mathcal{E}^2 - \mathcal{A}\mathcal{F}^2 + 2\mathcal{D}\mathcal{E}\mathcal{F} \\ &= \frac{1}{\sigma_f^2} \cdot \left[\sum_{ij} \frac{x_i^2 - x_i x_j}{\sigma_i^2 \sigma_j^2} + \sigma_f^2 \cdot \sum_{ijk} \frac{(x_i^2 - x_i x_j)y_k^2 - x_i^2 y_j y_k - x_i y_i x_j y_j + 2x_i y_i x_j y_k}{\sigma_i^2 \sigma_j^2 \sigma_k^2} \right] . \end{aligned} \quad (18)$$

By virtue of Eqs. (13) and (18), $S \cdot D_{\mathbf{T}} = \sigma_f^2 \cdot D_{\mathbf{\Lambda}}$, then comparing Eqs. (11) and (12) with Eqs. (16) and (17), the exactly same analytical results of corresponding quantities of two methods show their equivalence directly.

2.3 Discussion

If we denote $f_i(\vec{r})$ as a function at i -th point of any measurements $\vec{r} = \vec{r}(x, y, z, \dots)$ with uncertainty σ_i , the weighted average of $f_i(\vec{r})$ can be defined as

$$\bar{f}(\vec{r}) = \sigma_s^2 \cdot \sum_i \frac{f_i(\vec{r})}{\sigma_i^2} ,$$

with

$$\frac{1}{\sigma_s^2} \equiv \sum_{i=1}^n \frac{1}{\sigma_i^2} \quad \text{or} \quad \sigma_s^2 \equiv 1 / \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right).$$

Using above definition, we write the optimized quantities of Eqs. (16) and (17) in another forms, for example, we have

$$\hat{f} = \left(\frac{1}{\sigma_s^2} \cdot \overline{x^2} - \frac{\bar{x}}{\sigma_s^2} \cdot \frac{\bar{x}}{\sigma_s^2} \right) / (\sigma_f^2 \cdot D_{\Lambda}) \quad , \quad (19)$$

and

$$\sigma_{\hat{f}}^2 = \sigma_f^2 \cdot \left(\frac{1}{\sigma_s^2} \cdot \overline{x^2} - \frac{\bar{x}}{\sigma_s^2} \cdot \frac{\bar{x}}{\sigma_s^2} \right) / (\sigma_f^2 \cdot D_{\Lambda}) \quad . \quad (20)$$

It is interesting to notice, from Eqs. (19) and (20), we obtain the following relation

$$\sigma_{\hat{f}}^2 = \hat{f} \cdot \sigma_f^2 \quad . \quad (21)$$

Furthermore, taking advantage of the forms as displayed in Eqs. (19) and (20), we can acquire the relation

$$\hat{f} \cdot \bar{y} = \hat{\beta} + \hat{\alpha} \cdot \bar{x} \quad . \quad (22)$$

At the same time, from Eq. (5), we have

$$\hat{k}_x = \hat{\beta} + \hat{\alpha} \cdot x \quad . \quad (23)$$

Here subscript denotes the dependence of \hat{k} on x . So combining with Eq. (22), we get

$$\hat{k}_{\bar{x}} = \hat{f} \cdot \bar{y} \quad . \quad (24)$$

Comparing with the results in Ref. [7], we find the relations expressed in Eqs. (21) and (24) hold for both the constant and the linear fitting cases.

3 Experimental test

For the purpose of illustration, the application of the theoretical formulas to a simplified R value measurement is discussed in this section.

R , the ratio of the hadron production cross section via single photon annihilation to the lowest order point-like QED $\mu^+\mu^-$ cross section $\sigma_{pt} = 4\pi\alpha^2/3s$, is a fundamental quantity in e^+e^- interaction. It is calculated in the naive quark-parton model as $R = 3 \sum_q Q_q^2$, where Q_q is the quark electric charge, and the summation runs over all the produced flavors. Taking the lowest order QCD correction and the electro-weak effect into consideration, R value would be larger than the naive value (10/3), and the corrected term is a slowly varying smooth function of center-of-mass (C.M.) energy, in the region

Table 1: Values for R ^[9]. The errors quoted are point-to-point systematic errors.

E_{cm} (GeV)	R value	Error ΔR	E_{cm} (GeV)	R value	Error ΔR
5.60	4.08	0.32	6.60	4.50	0.17
5.70	4.09	0.16	6.65	4.25	0.16
5.75	4.12	0.20	6.70	4.63	0.15
5.80	4.13	0.16	6.75	4.38	0.15
5.85	4.13	0.19	6.80	4.44	0.16
5.90	4.09	0.14	6.85	4.50	0.13
5.95	4.17	0.16	6.90	4.41	0.15
6.00	4.17	0.09	6.95	4.23	0.17
6.05	4.16	0.18	7.00	4.10	0.12
6.10	4.04	0.15	7.05	4.31	0.09
6.15	4.34	0.16	7.10	4.32	0.14
6.20	4.05	0.08	7.15	4.29	0.11
6.25	3.96	0.14	7.20	4.27	0.11
6.30	4.27	0.14	7.25	4.39	0.11
6.35	4.47	0.17	7.30	4.29	0.11
6.40	4.31	0.13	7.35	4.33	0.09
6.45	4.23	0.14	7.40	4.46	0.08
6.50	4.40	0.15	7.45	4.51	0.14
6.55	4.66	0.16	7.50	4.18	0.59

without any resonances, therefore, R could reasonably be depicted by a linear function in a good approximation.

In experiment, many factors should be considered in R value calculation⁴. As a pedagogical example, a comparatively concise R expression ^[4] is given here

$$R = \frac{(N - N_{bg})}{L\epsilon(1 + \delta) \cdot \sigma_{pt}} ,$$

where N is the number of multi-hadronic events detected, N_{bg} is the estimated number of background events, L is the integrated luminosity, $\epsilon(1 + \delta)$ is the acceptance for the multi-hadronic events with radiative effect included and $(1 + \delta)$ is the radiative correction factor due to higher order QED processes up to order α^3 . Table 1 lists thirty eight experiment R -values^[9]. From a study of data taken at different times at the same C.M.

⁴The R value measurement at BESII has been described in Refs. [10, 11], where the detailed calculation about experiment R value could be found.

energy, the estimated systematic point-to-point errors are given as $\pm 3\%$. For the R value used here, the systematic uncertainty in the detection efficiency ($\pm 8\%$), the luminosity measurement ($\pm 6\%$), the event selection procedure ($\pm 2\%$), and the background subtraction ($\pm 3\%$) yielded a common systematic error of $\pm 10\%$, which should be considered as the normalization error. Now these thirty eight R -values will be used to test foregoing conclusions. For minimization, the MINUIT package, one of useful CERN packages in high energy physics ^[2], is utilized.

In the χ^2 construction, the following substitutes are adopted

$$x_i \rightarrow E_{cm}^i, \quad y_i \rightarrow R_{exp.}^i, \quad \sigma_i \rightarrow \Delta R_{exp.}^i, \quad \beta \rightarrow R_0, \quad \text{and} \quad \alpha \rightarrow \eta,$$

then Eqs. (7) and (6) become

$$\chi_f^2 = \sum_i \frac{[f R_{exp.}^i - (R_0 + \eta E_{cm}^i)]^2}{(\Delta R_{exp.}^i)^2} + \frac{(f - 1)^2}{\sigma_f^2},$$

$$\chi_M^2 = \sum_{ij} [R_{exp.}^i - (R_0 + \eta E_{cm}^i)] (\mathbf{V}^{-1})_{ij} [R_{exp.}^j - (R_0 + \eta E_{cm}^j)],$$

where σ_f is the overall error of normalization factor f , which equals to 10%, the element (v_{ij}) of matrix \mathbf{V} reads

$$v_{ij} = \delta_{ij} \cdot (\Delta R_{exp.}^i)^2 + \sigma_f^2 \cdot R_{exp.}^i \cdot R_{exp.}^j.$$

The fitting results are summarized in Table 2. At the same time, using Eqs. (16) and (17), we can compute the corresponding values theoretically, which are also given in Table 2. We can see different method leads to the same results up to the significant digits listed in the table. In addition, with these values, we can also test the simple relation given in Eq. (21).

Table 2: Experimental fitted and theoretical calculated values of parameters and relevant information.

Parameter	Matrix method	Factor method	Theoretical calculation
R_0	2.2895 ± 0.3772	2.2895 ± 0.3772	2.2895 ± 0.3772
η	0.1241 ± 0.0421	0.1241 ± 0.0421	0.1241 ± 0.0421
f	—	0.7282 ± 0.0853	0.7282 ± 0.0853
$\chi^2/\text{d.o.f}$	27.18/35	27.18/35	—

The fitted values relevant to R_0 and η show the minimization equivalence of two χ^2 forms fairly well. Next we turn to another aspect of two χ^2 forms, that is the fit biasness, which has been noticed in previous papers ^[5, 7, 8].

Fig. 1 shows the fitting result, where the solid line represents the best fitted k value. It is rather obvious that the fitted line is far below all data points. Merely by matrix χ^2 fit, we are at a loss what to do with this deviation. On the contrary, the factor χ^2 fit provides us a normalization factor f which just manifests the magnitude of the biasness. In fact, what we want to know is the weighted average of experiment data, that is \bar{y} , which should not contain the biasness due to common uncertainty. Eq. (24) gives the relation between \hat{k} and \bar{y} , from which we can obtain \bar{y} by scaling \hat{k} with the factor f . The dashed line in Fig. 1 denotes the factor scaled k value or expected \bar{y} . We can see the re-scaled line fits the experiment data fairly well⁵. After rescaled, $R_0/f = 3.14$ which is fairly consistent with the naive expectation $10/3 = 3.3$.

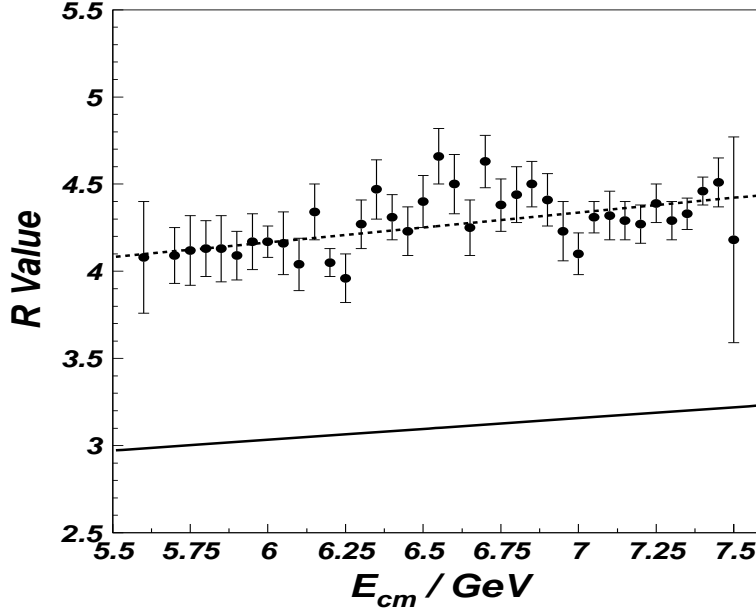


Figure 1: The R value, error bars indicate point-to-point systematic errors. The data points taken from Ref. [4]. The solid line represents the best fitted k value, while the dashed line the factor scaled k value or expected \bar{y} .

4 Summary

For linear function fitting, two χ^2 forms have been constructed to handle correlated data. The equivalence of these two forms has been proved strictly by mathematical calculation, and tested quantitatively by a typical simplified R value measurement experiment. However, in light of the comparison of two χ^2 forms, we notice that the factorized χ^2 form is more transparent, besides the concise expression and comparative easy calculation in

⁵In a forthcoming paper [12], we will examine the biasness of χ_f^2 for linear function fit in more detail.

minimization, the cause of the biasness is more easily recognized and corrected by the scale factor. By contrast, for matrix χ^2 form the biasness is uncontrolled.

The importance of this work lies in two aspects, first, the equivalence conclusion has been extended from constant fitting to linear function fitting; second, the analytical forms of the minimized parameters are worked out explicitly, which display some qualitatively features about more complex fitting.

At last, a remark is in order. Comparing with the formulas of the constant fitting, the formulas presented in this paper are more complicated and less intuitive, which implies that further extension of equivalence conclusion might be much difficult, or another approach should be found for equivalence study.

Appendix: Matrix Formulas

Some necessary matrix formulas are collected in this appendix, as to the knowledge of matrix, see Ref. [13].

Let's consider a special square matrix \mathbf{A} ,

$$\mathbf{A} = \begin{pmatrix} x_1 + \delta & a_1 b_2 + \delta & \cdots & a_1 b_n + \delta \\ a_2 b_1 + \delta & x_2 + \delta & \cdots & a_2 b_n + \delta \\ \vdots & \vdots & \ddots & \vdots \\ a_n b_1 + \delta & a_n b_2 + \delta & \cdots & x_n + \delta \end{pmatrix}. \quad (\text{A. 1})$$

If its matrix element reads

$$a_{ij} = \delta_{ij} \cdot (x_i - a_i b_i) + a_i b_j + \delta, \quad ,$$

then its element of the inverse matrix could be worked out

$$a_{ij}^{-1} = \frac{\delta_{ij}}{x_i - a_i b_i} - \frac{a_i b_j + \delta}{S \cdot (x_i - a_i b_i) \cdot (x_j - a_j b_j)} - \frac{\delta}{S} \cdot \sum_{k=1}^n \frac{(a_i - a_k) \cdot (b_j - b_k)}{(x_i - a_i b_i) \cdot (x_j - a_j b_j) \cdot (x_k - a_k b_k)},$$

with

$$S = 1 + \sum_{i=1}^n \frac{a_i b_i + \delta}{x_i - a_i b_i} + \delta \cdot \sum_{i=1}^n \sum_{j=1}^n \frac{a_i b_i - a_i b_j}{(x_i - a_i b_i) \cdot (x_j - a_j b_j)}.$$

In addition, its determinant has also been proved to be

$$\begin{vmatrix} x_1 + \delta & a_1 b_2 + \delta & \cdots & a_1 b_n + \delta \\ a_2 b_1 + \delta & x_2 + \delta & \cdots & a_2 b_n + \delta \\ \vdots & \vdots & \ddots & \vdots \\ a_n b_1 + \delta & a_n b_2 + \delta & \cdots & x_n + \delta \end{vmatrix} = S \cdot \left[\prod_{i=1}^n (x_i - a_i b_i) \right].$$

In fact, the error matrix given in equation (1) is just a special form of that given in equation (A. 1). So with above formulas, it is easy to acquire the corresponding results of matrix \mathbf{V} . If the element of \mathbf{V} reads

$$v_{ij} = \delta_{ij}\sigma_i^2 + y_i y_j \sigma_f^2 \quad ,$$

then the element of its inverse matrix is

$$\lambda_{ij} = v_{ij}^{-1} = \frac{\delta_{ij}}{\sigma_i^2} - \frac{\sigma_f^2}{S} \cdot \frac{y_i y_j}{\sigma_i^2 \sigma_j^2} \quad , \quad (\text{A. 2})$$

with

$$S = 1 + \sigma_f^2 \cdot \sum_{i=1}^n \frac{y_i^2}{\sigma_i^2} \quad . \quad (\text{A. 3})$$

References

- [1] CELLO Collab., H.-J.Behrend et al., Phys. Lett. **B182** (1987) 400-411.
- [2] CERN Library: “MINUIT Reference Manual” version 92.1(March 1992).
- [3] TASSO Collab., R.brandelik et al., Phys. Lett. **B113** (1982) 499-508;
TASSO Collab., R.brandelik et al., Z. Phys. **C4** (1980) 87-93.
- [4] JADE Collab., W.Bartel et al., Phys. Lett. **B129** (1983) 145-152.
- [5] G.D’Agostini, Nucl. Instr. Meth. **A 346** (1994) 306-311.
- [6] T.Takeuchi, Prog. Thero. Phys. Suppl. **123** (1996) 247-264.
- [7] Mo Xiao-Hu and Zhu Yong-Sheng, HEP & NP **27**, 371-376, (2003).
- [8] Mo Xiao-Hu and Zhu Yong-Sheng, HEP & NP **27**, 747-753, (2003).
- [9] J. L. Siegrist, et al., Phys. Rev. **D26**, 969-990 (1982).
- [10] J.Z.Bai et al., Phys. Rev. Lett. **84**, (2000) 594.
- [11] J.Z.Bai et al., Phys. Rev. Lett. **88**, (2002) 101802.
- [12] Mo Xiao-Hu, “Unbiased χ^2 Estimator for Linear Function Fit Involving Correlated Data”, paper in preparing.
- [13] Wener Greub, “Linear Algebra”(Fourth Edition),1981, Springer-Verlag New York Inc.